either the sense or antisense direction (relative to RgNanH). The recombination event was confirmed by PCR and the expression of the full cluster tested by qPCR. The expression of the genes flanking the cluster RUMGNA_02690 and 02702 showed levels comparable to the wild-type strain, as also observed for the first three genes of the nan cluster, RUMGNA_02701-02699, however, the nan cluster genes RUMGNA_02698-02691 showed significantly reduced expression compared to the wild-type strain.

[0110] To assess the effect of the nan cluster on the ability of *R. gnavus* to utilise sialic acid and sialoconjugates in vitro, *R. gnavus* ATCC 29149 wild-type and mutant strains were grown anaerobically with 3'SL or 2,7-anhydro-Neu5Ac. *R. gnavus* wild-type strain was able to utilise both 3'SL and 2,7-anhydro-Neu5Ac as a sole carbon source, but no growth was detected using the nan deletion mutants on these substrates (FIG. **9**), demonstrating the importance of the nan cluster to support growth of *R. gnavus* ATCC 29149 on these sialic acid derivatives.

In Vivo Colonisation of Germ-Free Mice by *R. gnavus* Wild-Type and Nan Mutants

[0111] To assess the impact of the nan cluster on the fitness of *R. gnavus* in vivo, germ-free C57BL/6J mice were gavaged with 1×10^8 CFU *R. gnavus* ATCC 29149 or *R. gnavus* antisense nan deletion mutant or a mixture of wild-type and nan mutant strains at 1×10^8 CFU each (FIG. **10**). During mono-colonisation experiments, both strains were detectable in the faecal content at day 3, 7 and 14 post-gavage at mean levels of between 1×10^6 and 1×10^7 bacteria per mg of material (FIG. **10***a*). Both strains were also detected in the caecal content of mono-colonised mice sacrificed at day 14. The absence of the nan cluster did not affect the mouse expression response, as shown by RNA seq. In competition experiments, primers based on the insertion in the RgNanH gene were used to distinguish between wild-type and nan mutant, the wild-type strain reached mean colonisation levels comparable to the levels obtained during mono-colonisation, whereas the mutant strain was severely outcompeted, reaching only 2×10^4 copies per mg at day 3, before decreasing further at day 7 and day 14 to levels below the level of detection, in both the faecal and caecal contents (FIG. **10***b*)

[0112] The impact of the nan deletion on the location of *R. gnavus* within the mucus layer was determined in mono-colonised mice by measuring the distance of the nan mutant or wild-type *R. gnavus* strains to the epithelial layer throughout the colon by fluorescent in situ hybridization (FISH) staining using confocal microscopy. The data showed that the nan mutant resided 19.70 μm from the epithelial layer, 5.06 μm further away than the wild-type strain, 14.64 μm (FIG. **10***c*&*d*).

Bioinformatics Search for Predicted Homologous Nan Clusters

[0113] Since the success of the *R. gnavus* niche competition strategy depends on the organism's ability to exclusively utilize 2,7-anhydro-Neu5Ac, we searched the database for predicted homologous nan clusters to estimate how widely distributed this strategy is among bacterial isolates. MultiGeneBlast analysis revealed that predicted homologs of the *R. gnavus* nan cluster are shared by a limited number of species, including 37 homologous clusters in *Streptococcus pneumoniae* isolates (illustrated in FIG. **11** by the functionally characterized NanB from *S. pneumoniae* D39, Manco et al., 2006), *S. suis* A7, *Blautia hansenii* DSM 20583, *Blautia* sp. YL58 and *Intestinimonas butyriciproducens* AF211 (FIG. **11**). This is also in line with the SSN bioinformatics analysis reported in FIG. **5**, showing a range of species encoding NanC or IT-sialidase like genes.

[0114] In addition to the presence of a predicted IT-sialidase, the clusters share a predicted ROK family kinase, oxidoreductase, β-galactosidase, Neu5Ac lyase, and ManNAc-6-P epimerase (FIG. **11**). All **37** *S. pneumoniae* NanB clusters share a similar organization and the more variable area between the two subclusters (white in FIG. **11**) contains an additional ABC transporter compared to the other nan clusters. These *Streptococcus* clusters harbour a RpiR-type regulator (pink), whereas an AraC-type regulator (purple) is present in the nan clusters of the other bacterial species. *Blautia* sp. YL58 has the only nan cluster that contains a RUMGNA_RS11885 lipase/esterase homolog (grey), yet both the *S. suis* A7 and *I. butyriciproducens* AF211 clusters contain a different type of esterase (yellow).

[0115] A major difference between NanB/NanH IT-sialidase and NanC sialidase cluster types is the associated transporter class, a carbohydrate ABC transporter for NanB/NanH (jade green) as opposed to a sodium:solute symporter in NanC clusters (Xu et al., 2011), which may indicate a difference in the form of sialic acid being transported. Altogether, these analyses support the specialisation of the *R. gnavus* nan cluster, conferring the bacteria with a unique advantage over other members of the gut microbiota to colonise the mucus niche in the human colon.

1. A method of identifying, monitoring and/or diagnosing mucosal bacterial presence or infection, said method including the step of detecting at least part of a sialic acid transporter protein encoded by *Ruminococcus gnavus* (*R. gnavus*) ATCC 29149 Nan cluster.

2. A method according to claim **1** wherein the transporter protein is specific to 2,7-anhydro-Neu5Ac.

3. A method according to claim **1** wherein the substrate or solute binding protein of the ATCC 29149 Nan cluster is encoded by RUMGNA_02698.

4. A method according to claim **1** wherein the transporter protein is used as an indicator or biomarker for inflammatory bowel disease.

5. A method according to claim **4** wherein the transporter protein is used as a faecal biomarker.

6. A method according to claim **1** wherein the presence of the transporter protein is used as an indicator of likelihood of success of microbiome-targeted therapies

7. A method according to claim **6** wherein the therapy is faecal microbiota transplantation.

8. A method according to claim **1** wherein polymerase chain reaction (PCR) is used to amplify the protein and/or identify the presence of the transporter protein.

9. A method according to claim **8** wherein quantitative polymerase chain reaction (qPCR) is used to identify the presence of the transporter protein.

10. A method according to claim **1** wherein the presence or absence of the transporter protein is used to distinguish or diagnose Ulcerative Colitis or Crohn's Disease.

11. A method of inhibition of the growth of bacterium, said method including the step of inhibition of a sialic acid transporter protein.